

# SEQUENCE LISTING

<110> Black Jr., Charles A.

<120> COMPOSITIONS AND METHODS FOR ACTIVATING  
GENES OF INTEREST

<130> 5722-2(35722/191928)

<140> 09/446,402

<141> 1999-12-20

<150> PCT/US98/13093

<151> 1998-06-24

<150> 60/050,772

<151> 1997-06-25

<160> 19

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 4279

<212> DNA

<213> Artificial Sequence

<220>

<223> Recombinant Molecule containing multiple cloning  
site, kozak sequence, LacZ gene.

<221> misc\_feature

<222> (1)...(64)

<223> Multiple cloning site

<221> misc\_feature

<222> (65)...(79)

<223> Consensus sequence for the "Kozak sequence"  
(translation initiation)

<221> prim\_transcript

<222> (80)...(4279)

<223> Beta galactosidase

<400> 1

|            |            |            |            |            |             |     |
|------------|------------|------------|------------|------------|-------------|-----|
| ttaatacgac | tcactatagg | ctagcctcga | gaattcacgc | gtggtacctc | tagagtcgac  | 60  |
| ccgggccgcc | gccaccatgg | cgcagcacca | tggcctgaaa | taacctctga | aagaggaact  | 120 |
| tggttaggta | ccttctgagg | cggaaagaac | cagctgtgga | atgtgtgtca | gttaggggtg  | 180 |
| ggaaagtccc | caggctcccc | agcaggcgca | agtatgcaaa | gcatgcatct | caattagtca  | 240 |
| gcaaccaggt | gtggaaagtc | cccaggctcc | ccagcaggca | gaagtatgca | aagcatgcat  | 300 |
| ctcaattagt | cagcaaccat | agtcccgcgc | ctaactccgc | ccatcccgcc | cctaactccg  | 360 |
| cccagttccg | cccattctcc | gccccatggc | tgactaattt | tttttattta | tgacagaggcc | 420 |
| gaggccgcct | cggcctctga | gctattccag | aagtagtgag | gaggcttttt | tggaggccta  | 480 |
| ggcttttgca | aaaagcttgg | gatctctata | atctcgcgca | acctattttc | ccctcgaaca  | 540 |

|             |             |             |             |             |            |      |
|-------------|-------------|-------------|-------------|-------------|------------|------|
| ctttttaagc  | cgtagataaa  | caggctggga  | cacttcacat  | gagcgaaaaa  | tacatcgta  | 600  |
| cctgggacat  | gttgacagatc | catgcacgta  | aactcgcaag  | ccgactgatg  | ccttctgaac | 660  |
| aatggaaagg  | cattattgcc  | gtaagccgtg  | gcggtctggt  | accggtgggt  | gaagaccaga | 720  |
| aacagcacct  | cgaactgagc  | cgcgatattg  | cccagcgttt  | caacgcgctg  | tatggcgaga | 780  |
| tcgatcccgt  | cgttttacaa  | cgctcgtgact | gggaaaaccc  | tggcgttacc  | caacttaac  | 840  |
| gccttgacgc  | acatccccct  | ttcgccagct  | ggcgtaatag  | cgaagaggcc  | cgcaccgatc | 900  |
| gcccttccca  | acagttgcgc  | agcctgaatg  | gcgaatggcg  | ctttgcctgg  | tttccggcac | 960  |
| cagaagcggg  | gccggaagc   | tggctggagt  | gcgatcttcc  | tgaggccgat  | actgtcgctg | 1020 |
| tcccctcaaa  | ctggcagatg  | cacggttacg  | atgcgcccat  | ctacaccaac  | gtaacctatc | 1080 |
| ccattacggg  | caatccgcgc  | tttgttccca  | cggagaatcc  | gacgggttgt  | tactcgctca | 1140 |
| catttaatgt  | tgatgaaagc  | tggctacagg  | aaggccagac  | gcgaattatt  | tttgatggcg | 1200 |
| ttactcggc   | gtttcatctg  | tgggtcaacg  | ggcgctgggt  | cggttacggc  | caggacagtc | 1260 |
| gtttgcgctc  | tgaatttgac  | ctgagcgcac  | ttttacgcgc  | cggagaaaac  | cgcctcgcg  | 1320 |
| tgatggtgct  | gcgttggagt  | gacggcagtt  | atctggaaga  | tcaggatatg  | tggcggatga | 1380 |
| gcggcatttt  | ccgtgacgct  | tcgttgctgc  | ataaacgcac  | tacacaaatc  | agcgatttcc | 1440 |
| atgttgccac  | tcgctttaat  | gatgatttca  | gccgcgctgt  | actggagggt  | gaagttcaga | 1500 |
| tgtgcggcga  | gttgcggtgac | tacctacggg  | taacagtttc  | tttatggcag  | ggtgaaacgc | 1560 |
| aggctcgccag | cggcaccgcg  | cctttcggcg  | gtgaaattat  | cgatgagcgt  | ggtggttatg | 1620 |
| ccgatcgcgt  | cacactacgt  | ctgaacgctg  | aaaaccgaa   | actgtggagc  | gccgaaatcc | 1680 |
| cgaatctcta  | tcgtgcgggtg | gttgaactgc  | acaccgccga  | cggcacgctg  | attgaagcag | 1740 |
| aagcctgcga  | tgtcgggtttc | cgcgaggtgc  | ggattgaaaa  | tggctctgctg | ctgctgaacg | 1800 |
| gcaagccgtt  | gctgattcga  | ggcgtaaac   | gtcacgagca  | tcatectctg  | catggtcagg | 1860 |
| tcattgatga  | gcagacgatg  | gtgcaggata  | tcctgctgat  | gaagcagaac  | aactttaacg | 1920 |
| ccgtgcgctg  | ttcgcatatt  | ccgaaccatc  | cgtgtgtgta  | cacgctgtgc  | gaccgctacg | 1980 |
| gacctgatgt  | ggtggatgaa  | gccaatattg  | aaaccacgg   | catggtgcca  | atgaatcgct | 2040 |
| tgaccgatga  | tcgcgcgtgg  | ctaccggcga  | tgagcgaacg  | cgtaacgcga  | atggtgcagc | 2100 |
| gcgatcgtaa  | tcacccgagt  | gtgatcatct  | ggtcgctggg  | gaatgaatca  | ggccacggcg | 2160 |
| ctaatacaga  | cgcgctgtat  | cgtggatca   | aatctgtcga  | tccttcccgc  | ccggtgcagt | 2220 |
| atgaaggcgg  | cggagccgac  | accacggcca  | ccgatattat  | ttgcccgatg  | tacgcgcgcg | 2280 |
| tggatgaaga  | ccagcccttc  | ccggctgtgc  | cgaaatggtc  | catcaaaaaa  | tggctttcgc | 2340 |
| tacctggaga  | gacgcgcccc  | ctgatecttt  | gcgaatacgc  | ccacgcgatg  | ggtaacagtc | 2400 |
| ttggcggttt  | cgctaaatac  | tggcaggcgt  | ttcgtcagta  | tccccgttta  | cagggcggt  | 2460 |
| tcgtctggga  | ctgggtggat  | cagtcgctga  | ttaaataatga | tgaaaacggc  | aacccggtgt | 2520 |
| cggcttacgg  | cgggtgatttt | ggcgatacgc  | cgaacgatcg  | ccagttctgt  | atgaacggtc | 2580 |
| tggctcttgc  | cgaccgcacg  | ccgcatccag  | cgctgacgga  | agcaaaaacac | cagcagcagt | 2640 |
| ttttccagtt  | ccgtttatcc  | gggcaaacca  | tcgaagtgc   | cagcgaatac  | ctgttccgct | 2700 |
| atagcgataa  | cgagctcctg  | cactggatgg  | tggcgctgga  | tggtaagccg  | ctggcaagcg | 2760 |
| gtgaagtgcc  | tctggatgtc  | gctccacaag  | gtaaacagtt  | gattgaactg  | cctgaactac | 2820 |
| cgcagccgga  | gagcgccggg  | caactctggc  | tcacagtacg  | cgtagtgcaa  | ccgaacgcga | 2880 |
| ccgcatggtc  | agaagccggg  | cacatcagcg  | cctggcagca  | gtggcgctctg | gcggaaaacc | 2940 |
| tcagtgtgac  | gctccccgcc  | gcgtcccacg  | ccatcccga   | tctgaccacc  | agcgaaatgg | 3000 |
| atttttgcat  | cgagctgggt  | aataagcgtt  | ggcaatttaa  | ccgccagtca  | ggctttcttt | 3060 |
| cacagatgtg  | gattggcgat  | aaaaaacaac  | tgctgacgcc  | gctgcgcgat  | cagttcaccc | 3120 |
| gtgcaccgct  | ggataacgac  | attggcgtaa  | gtgaagcgac  | ccgcattgac  | cctaacgcct | 3180 |
| gggtcgaacg  | ctggaaggcg  | gcgggccatt  | accaggccga  | agcagcggtg  | ttgcagtgca | 3240 |
| cggcagatac  | acttgctgat  | gcggtgctga  | ttacgaccgc  | tcacgcgtgg  | cagcatcagg | 3300 |
| ggaaaacctt  | atztatcagc  | cggaaaacct  | accggattga  | tggtagtggg  | caaattggcg | 3360 |
| ttaccgttga  | tgttgaaagt  | gcgagcgata  | caccgcaccc  | ggcgcggtat  | ggcctgaact | 3420 |
| gccagctggc  | gcaggtagca  | gagcgggtaa  | actggctcgg  | attagggccg  | caagaaaact | 3480 |
| atcccagacc  | ccttactgcc  | gcctgttttg  | accgctggga  | tctgccattg  | tcagacatgt | 3540 |
| ataccccgta  | cgtcttcccg  | agcgaaaacg  | gtctgcgctg  | cgggacgcgc  | gaattgaatt | 3600 |
| atggcccaca  | ccagtggcgc  | ggcgacttcc  | agttcaacat  | cagccgctac  | agtcaacagc | 3660 |
| aactgatgga  | aaccagccat  | cgccatctgc  | tgcacgcgga  | agaaggcaca  | tggctgaata | 3720 |
| tcgacgggtt  | ccatatgggg  | attggtggcg  | acgactcctg  | gagcccgtca  | gtatcggcgg | 3780 |
| aattccagct  | gagcgccggg  | cgctaccatt  | accagttggg  | ctggtgtcaa  | aaataataat | 3840 |

|            |            |            |             |            |            |      |
|------------|------------|------------|-------------|------------|------------|------|
| aaccgggag  | gccatgtctg | cccgtatttc | gcgtaaggaa  | atccattatg | tactatttaa | 3900 |
| aaaacacaaa | cttttggatg | ttcggtttat | tctttttctt  | ttactttttt | atcatgggag | 3960 |
| cctacttccc | gtttttcccg | atttggctac | atgacatcaa  | ccatatcagc | aaaagtgata | 4020 |
| cgggtattat | ttttgccgct | atttctctgt | tctcgctatt  | attccaaccg | ctgtttggtc | 4080 |
| tgctttctga | caaactcgga | acttgtttat | tcgagcttat  | aatggttaca | aataaagcaa | 4140 |
| tagcatcaca | aatttcacaa | ataaagcatt | tttttctactg | cattctagtt | gtggtttgtc | 4200 |
| caaactcatc | aatgtatctt | atcatgtctg | gacccctctag | agtcgacctg | caggcatgca | 4260 |
| agctggcact | ggccgctcgt |            |             |            |            | 4279 |

<210> 2  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic oligonucleotide

|                        |    |
|------------------------|----|
| <400> 2                |    |
| gaatacaaaag cttatgcatg | 20 |

<210> 3  
 <211> 13  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic oligonucleotide

|                 |    |
|-----------------|----|
| <400> 3         |    |
| gaatacaaaag ctt | 13 |

<210> 4  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic oligonucleotide

|                       |    |
|-----------------------|----|
| <400> 4               |    |
| aaagcttatg catgcggccg | 20 |

<210> 5  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic oligonucleotide

|                       |    |
|-----------------------|----|
| <400> 5               |    |
| cggccgcatc tagagggccc | 20 |

<210> 6  
 <211> 25

|                                  |    |
|----------------------------------|----|
| <212> DNA                        |    |
| <213> Artificial Sequence        |    |
| <220>                            |    |
| <223> Synthetic oligonucleotide  |    |
| <400> 6                          | 25 |
| gcggccgcat ctagagggcc cggat      |    |
| <210> 7                          |    |
| <211> 24                         |    |
| <212> DNA                        |    |
| <213> Artificial Sequence        |    |
| <220>                            |    |
| <223> Synthetic oligonucleotide  |    |
| <400> 7                          | 24 |
| aatacaaagc ttatgcatgc ggcc       |    |
| <210> 8                          |    |
| <211> 30                         |    |
| <212> DNA                        |    |
| <213> Artificial Sequence        |    |
| <220>                            |    |
| <223> Synthetic oligonucleotide  |    |
| <400> 8                          | 30 |
| aatacaaagc ttatgcatgc ggccgcatct |    |
| <210> 9                          |    |
| <211> 20                         |    |
| <212> DNA                        |    |
| <213> Artificial Sequence        |    |
| <220>                            |    |
| <223> Synthetic oligonucleotide  |    |
| <400> 9                          | 20 |
| catgcataag ctttgtattc            |    |
| <210> 10                         |    |
| <211> 13                         |    |
| <212> DNA                        |    |
| <213> Artificial Sequence        |    |
| <220>                            |    |
| <223> Synthetic oligonucleotide  |    |
| <400> 10                         | 13 |
| aagctttgta ttc                   |    |
| <210> 11                         |    |
| <211> 20                         |    |

100260-09560-95660

<212> DNA  
<213> Artificial Sequence

<220>  
<223> Synthetic oligonucleotide

<400> 11  
cggccgcatg cataagcttt 20

<210> 12  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Synthetic oligonucleotide

<400> 12  
gggccctcta gatgcggccg 20

<210> 13  
<211> 25  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Synthetic oligonucleotide

<400> 13  
atccggggccc tctagatgcg gccgc 25

<210> 14  
<211> 24  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Synthetic oligonucleotide

<400> 14  
ggccgcatgc ataagctttg tatt 24

<210> 15  
<211> 30  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Synthetic oligonucleotide

<400> 15  
agatgcggcc gcatgcataa gctttgtatt 30

<210> 16  
<211> 1798

<212> RNA  
<213> Unknown

<220>  
<223> mRNA sequence for Firefly luciferase

<400> 16  
gaauacaaag cuuaugcaug cggccgcauc uagagggccc ggauccaaaau ggaagacgcc 60  
aaaaacauaa agaaaggccc ggcgccauuc uauccucuag aggauggaac cgcuggagag 120  
caacugcaua aggcuauaag gagauacgcc cugguuccug gaacaauugc uuuuacagau 180  
gcacauaucg aggugaacau cacguacgcg gaauacuucg aaauuguccg ucgguuggca 240  
gaagcuaua aacgauaugg gcugaauaca aauacagaa ucgucguaug cagugaaaac 300  
ucucuuaau ucuuuauugc gguguugggc gccguuuuu aucggaguug caguugcgcc 360  
cgcgaaagcac auuuauaau aacgugaauu gcucaacagu augaacaauu cgcagccuac 420  
cguaguguuu guuuccaaaa agggguugca aaaaauuug aacgugcaaa aaaaauuacc 480  
aauaauccag aaaaauuua ucauggauuc uaaaacggau uaccagggau uucagucgau 540  
guacacguuc gucacauuc aucuaccucc cgguuuuuu gaauacgauu uuguaccaga 600  
guccuuugau cgugacaaaa caauugcacu gauaauaau uccucuggau cuacuggguu 660  
accuaagggg ugggcccuuc cgcauagaac ugccugcguc agauucucgc augccagaga 720  
uccuauuuuu ggcaaucaaa ucauuccgga uacugcgauu uuaagugug uuccauucca 780  
ucacgguuuu ggauguuua cuacacucgg auuuuugaua uguggauuuc gagucgucuu 840  
aauguauaga uuugaagaag agcuguuuuu acgaucuuu caggauuaca aaaucaaaag 900  
ugcgugucua guaccaaccc uauuuucauu cuucgccaaa agcacucuga uugacaaaau 960  
cgauuuaucu aaauuacacg aaauugcuuc uggggggcgca ccucuucga aagaagucgg 1020  
ggaagcgguu gcaaaacgcu uccaucuucc agggauacga caaggauaug ggcucacuga 1080  
gacuacauga gcuauucuga uuacaccga gggggauaug aaaccgggcg cggucgguaa 1140  
aguuguucca uuuuuugaag cgaagguugu ggaucuggau accgggaaaa cgcugggcg 1200  
uaaucagaga ggcgaauuau gugucagagg accuauaau auguccgguu auguaaaca 1260  
uccggaagcg accaacgcu ugaugacaa ggauggaug cuacauucg gagacauagc 1320  
uuacugggac gaagacgaac acuucuau aguugaccgc uugaagucuu uaauuaaaua 1380  
caaaggauau cagguggccc ccgcugaau ggaucgaua uuguuacaac accccaacau 1440  
cuucgacgcg ggcguggcag gucuuccga cgaugacgcc ggugaacuuc ccgccgccgu 1500  
uguuguuuug gagcacggaa agacgaugac ggaaaaagag aucguggauu acgucgccag 1560  
ucaaguaaca accgcgaaaa aguugcgcg aggaguugug uuuguggacg aaguaccgaa 1620  
aggucuuacc ggaaaacucg acgcaagaaa aaucagagag auccucaua aggccaagaa 1680  
ggcgcgaaag uccaaauugu aaaauguaac uguauucagc gaugacgaaa uucuuaagcu 1740  
uuguaauccu ccgagggggc gagcucccaa aaaaaaaaaa aaaaaaaaaa aaaaaaaa 1798

<210> 17  
<211> 12  
<212> DNA  
<213> Consensus Kozack sequence

<400> 17  
gccgccgcca tg 12

<210> 18  
<211> 13  
<212> DNA  
<213> Consensus Kozack sequence

<400> 18  
gccgccrcca ugg 13

<210> 19

**THE UNIVERSITY OF CHICAGO**

8